SYSTEMATIC AND STATISTICAL VARIATIONS OF OMMATIDIAL AREA IN THE DROSOPHILA EYE AND THEIR EFFECT ON EYE MORPHOLOGY

BY

BOYUAN YANG

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Theoretical and Applied Mechanics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Advisor:

Professor Sascha Hilgenfeldt

ABSTRACT

The study of metamorphosis and morphogenesis in the Drosophila fruit fly has far-reaching consequences for our fundamental understanding of these processes and wide-spread applications in biomedicine. Understanding the influence of mutations in the planar cell polarity pathway is one important aspect in this respect, and will potentially help in the treatment of diseases. The Drosophila eye consists of small units called ommatidia, shaped as regular hexagons under the balance of cell-cell adhesion and interfacial tension. Such regular hexagonal tessellations are known to obey certain topological and geometrical relations. In particular, the theory of Voronoi tilings shows that only a small magnitude of area variation is allowed before defects occur. Yet, the experiment image shows area variations far greater than this limit. A theory is proposed separating the variation into two independent parts, i.e., systematical and statistical variations. The statistical variation describes the local magnitude of disorder and should follow the variation restriction. The systematical variation describes area changes over large scales. The total variation observed in experiment is the combination of the two variations. This hypothesis is verified both in simulations and in analyzing experimental images of wild-type and mutant Drosophila retinas.

ACKNOWLEDGMENTS

I wish to express my appreciation to Professor Hilgenfeld for his guidance and patience during my research. I would never be able to feel the thrill of new research findings with out him. In addition, I want to thank Dr. Justin Cassidy and Professor Richard Carthew for providing the high quality images and unconditional assistance whenever needed. Finally, I want to thank my parents for supporting me during my graduate years in a foreign country thousands miles from home.

Table of Contents

LIST OF SYMBOLS

c_A	Coefficient of Variation of Area
$c_{A,sys}$	Coefficient of Systematical Variation of Area
$c_{A,stat}$	Coefficient of Statistical Variation of Area
$c'_{A,stat}$	Calculated Coefficient of Statistical Variation of Area
$c_{A,tot}$	Coefficient of Variation of Area in Experiment
c_n	Coefficient of Variation of Number of Neighbor
$\overline{n}$	Average Number of Neighbor
r_0	Grid Lateral Length
α	Magnitude Parameter of Gaussian Noise
β	Magnitude Parameter of Imposed Area Gradient
μ_A	Average Area
σ_n	Variation of Number of Neighbor

CHAPTER 1: INTRODUCTION

Morphogenesis is the biological process that causes cells to have various shapes and arrangements. Morphogenesis follows certain mechanical rules, which are controlled by gene expression. The genes, called orthologs, are believed to retain the same function across species [1]. Such preservation of gene function makes it possible to study morphogenesis in Drosophila (by far the most common genetic model organism) and expand fundamental and applied knowledge to important fields like disease treatment. Some biological tissues have extremely regular, ordered patterns. For instance, the Drosophila retina consists of regular hexagonal ommaditia [2] [3]. Genetic mutations that influence the morphological process will distort the regular pattern and form defects. The study of topological and geometrical properties of these patterns gives insight into the mechanics of morphogenesis.

1.1 Topological and Geometrical Properties

Packing is an important field in the study of cellular structure. Very often, the detailed structure of the packing is variable, but some topological and geometrical properties can describe the nature of the problem clearly in statistical terms. Meanwhile, 3-D packing problems are often reduced to 2-D planar tessellation problems if the changes along the third dimension are negligible. Such reduction will greatly reduce the difficulty of the problem and many existing theories can be applied.

1.1.1 Euler's Law

Euler's law is one of the most fundamental topological characters in 2-D planar tessellation. It can be written in one simple equation,

$$F + V - E = \chi \tag{1.1}$$

where F is the number of faces, V is the number of vertices and E is the number of edges. The Euler characteristic χ is an integer, which is dependent only on the structures topological genus.

In a planar cellular structure, a triple junction of edges is the only stable structure, if the edges are homogeneous. Four-way junctions are unstable, but can appear as an intermediate stage in processes like T-1 transition. The T-1 transition is a local neighbor change process, which is illustrated in Figure 1.1. Under such circumstances, each cell represents one face, each cell-cell interface represents one edge and each junction of three cells represents one vertex. Assuming each face F_i has n sides, we reaches the following set of equations,

$$\begin{cases} \sum F_i = F \\ \sum F_i \cdot n = 2 \cdot E \\ \sum F_i \cdot n = 3 \cdot V \end{cases} \tag{1.2}$$

Substituting the relation back into equation (1.1) result in the following,

$$\sum F_i \cdot \left(1 - \frac{1}{6}n\right) = \chi \tag{1.3}$$

The average number of faces is defined as $\overline{n} = \sum(n \cdot F_i)/\sum F_i$, which gives

$$\overline{n} = 6 - \frac{\chi}{F} \tag{1.4}$$

As the number of faces approaching infinity, the average number of neighbors reaches six.

The six-neighbor property is preserved for 2D tilings in various different fields, like biological structures (Drosophila wing epithelium, Drosophila eye), physical structures (2-D foam) and purely mathematical tessellations (Possion-Voronoi tiling) [2] [4] [5].

1.1.2 Area and Number of Neighbor Relation

Numerous studies build a connection between area, a geometrical property, and the number of neighbors, a topological property. In 1928, Lewis discovered an empirical linear dependence of the average areas of n-sided cells on the number of sides n [6]. Later, empirical probability distributions for the number of sides and areas were found as 2-parameter gamma distributions [5]; the distributions display further correlations such as Lemaitre's law [7]. Recently, an analytical result was derived for the relation between the neighbor probabilities and the coefficient of variation of area, using a local model focusing on a single domain (cell) and its nearest neighbors only [8]. These statistical relations provide a powerful tool for the understanding of cellular structures in industrial materials or biological tissues.

1.1.3 Symmetry Break in Hexagonal Voronoi Tiling

Lucarini studied the relation between the number of neighbors and area from a different angle [9]. Instead of a local model, he started with the Voronoi tiling of a large set of points in 2-D space. The Voronoi tiling in 2-D is a partition of the plane. Let X be a 2-D space and $P = \{p_1, p_2, \cdots, p_n\}$ is a subsets of X. A distance function d is defined in that space. The Voronoi cell V_i corresponding to point p_i is defined as $V_i = \{x \in X | d(x, P_i) \leq d(x, P_k), \forall k \neq i\}$. Choosing different distance function, the result will be different. In this case, the distance function is the Euclidean distance.

A Voronoi tiling can generate regular patterns. If the points follow an equal-lateral triangular grid with edge length r_0, then the corresponding Voronoi tiling will consist of regular hexagon, see Figure 1.2. Then, a perturbation is introduced by adding Gaussian noise to the spatial position of the points, with a magnitude

$|\varepsilon|^2 = \alpha^2 r_0^2$. Topological and geometrical properties are examined using a large sample size. Euler's law is checked. The average number of neighbor remains approximately six regardless of the magnitude of noise. The dependence of coefficient of variation of area $c_A = \sigma_A/\mu_A$, which is the standard deviation of area divided by mean area, and standard deviation of number of neighbors σ_n on the strength of noise is revealed, see Figure 1.3. The relation shows that, while c_A increases with α, all cells remain six-neighbor cells ($\sigma_n = 0$) if α stays below a critical value $\alpha_c \approx 0.12$. As α increases further, both c_A and σ_n will reach plateaus asymptotically. Lucarini also argues for a linear dependence $c_A \approx 0.583\alpha$ for small α [9]. It indicates that a reasonably small statistical variation of area is allowed for typical hexagonal tessellations, without topological defects. The critical value for the coefficient of variation of area is 0.07 as suggested by Lucarini. Meanwhile, the analytical result suggests that for defects to occur in a tessellation with roughly 700 cells, the critical value is $c_{A,critical} \approx 0.068$ [8]. It shows consistency in simulation and analytical result. For consistency purpose, the rest of the article will use coefficient of variation of number of neighbor $c_n = \sigma_n/\mu_n$ to measure the number of neighbor variation, which is equivalent to σ_n because the average number of neighbor remains six to very good approximation.

1.2 Mechanism in Drosophila Eye

Regular hexagonal pattern can be seen in many cases in nature, from honeycombs in a beehive to Rayleigh-Benard convection cells [10]. The Drosophila eye has a similar hexagonal pattern for wild-type flies [2] [3]. A typical drosophila eye consists of about 700 ommatidia in regular hexagonal shape. Each ommatidium has four cone cells in the center, two primary pigment cells enclose the cone cells and

six secondary and six tertiary cells at the perimeter, see Figure 1.4 and Figure 1.5. The formation of this cellular structure of ommatidia has been studied in detail [2] [11] [12] [13] [14]. It is believed that the mechanical force plays an important role in the ommatidia pattern formation. Cadherins within the adherens junction increase cell-cell adhesion. Combined with an effective interfacial tension, resulting from actin contractility and membrane tension, stable cellular structures are reached as an energetically favored state.

The morphogenesis of the retina is also studied genetically [2]. Cell-cell interaction and certain signaling pathways are believed to play a crucial role in the development of the retina from the unpatterned epithelium at early stages to the final complex heterogenous pattern with hexagonal order. Frizzled is a component of a signaling pathway that controls planar cell polarity [15] [16]. Frizzled mutations will affect morphogenesis at early stages and the cellular structure will be different from the regular case. Statistical correlation studies for the mutated pattern will give insight on the exact mechanism of signaling pathway and planar cell polarity. In the present work, we pursue the hypothesis that the failure of proper planar polarity in the mutants is not a direct consequence of the mutation, but of a greater variation of ommatidial area induced by the mutation, which then necessitates the observed orientational disorder.

1.3 Overview

As suggested by Lucarini, only a relatively small magnitude of area variation is permitted for the tessellation to remain hexagonal. In reality, we find that the area variation of ommatidia even in the wild type drosophila far exceeds the critical value, see Table 1.1. Normally, we will expect tessellation in nature follows the same area

variation restriction. So a bridge has to be built to link the theory and experimental result.

First, we define the overall variation measurements $c_{A,tot}$ and $c_{n,tot}$, which are calculated directly from sample raw data. As mentioned before, the experiment images indicate that $c_{A,tot}$ for wild drosophila ommatidia is on average 60% larger than the critical value 0.07.

The simulation performed by Lucarini has a uniform Gaussian noise applied to the whole sample. We denote the resultant variation as statistical variation. Further, two variables are introduced. First is the coefficient of statistical variation of area $c_{A,stat}$, which is the same as the coefficient of variation of area c_A introduced before and depends only on the strength of Gaussian noise α. . Similarly, we obtain a coefficient of statistical variation of the number of neighbors $c_{n,stat}$ due to Gaussian noise.

What is the cause of the additional area variation obtained from experiment? Observation of the experimental images suggests that there could be an overall area gradient on the scale of the entire eye. This will result in a systematical variation introduced by large-scale area gradient, in addition to the statistical variation. The ommatidium centroid describes its position in x-y plane. Using the empirical area values as a data set, we postulate a systematic area function $f(x, y)$ and use least squares fitting to obtain the parameters for this function, so that the systematical variation for overall area can be quantified. For instance, if the area for ommaditium decreases linearly along a certain direction, we can use a plane fitting and the coefficient of variation of area for the plane is the corresponding coefficient of systematical variation of area $c_{A,sys}$. The fitting is not limited to plane. Other fittings, like quadratic, might give better description of systematical variation. The fitting is

not limited to planes. Other fittings, like quadratic, might give better description of systematical variation. The procedure is described in detail in section 3.1.

The interaction of the two kinds of variation determines whether the tessellation remains hexagonal or has defects. Assuming that the systematical and statistical variations are independent with each other, the variations must add, i.e.,

$$c_{A,sys}^2 + c_{A,stat}^2 = c_{A,tot}^2 \tag{1.4}$$

With $c_{A,tot}$ from raw data and $c_{A,sys}$ from surface fitting, $c_{A,stat}$ can be calculated. For experiment images, the derived $c_{A,stat}$ is below the critical value for defect generation for wild image and above the critical value for the mutant image. Thus, the results are consistency with the previously presented theory, as the wild-type eyes have no defects, while the frizzled mutant eyes do show a number of prominent defects.

1.4 Figures

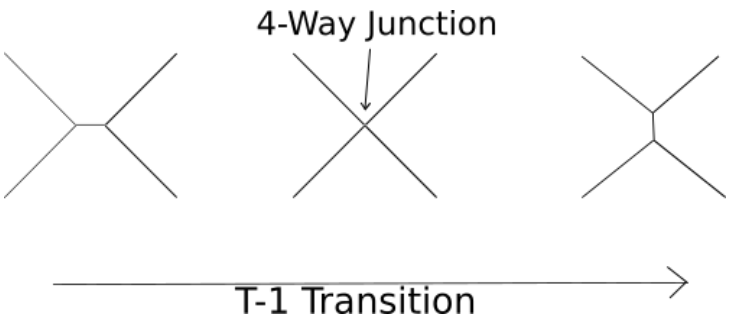

Figure 1.1: A typical T-1 transition process

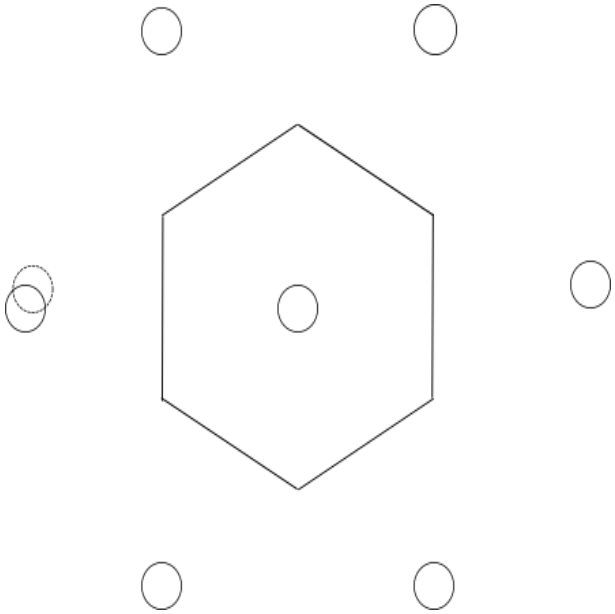

Figure 1.2: Voronoi tiling from triangular grid points with example of position variation on the left point

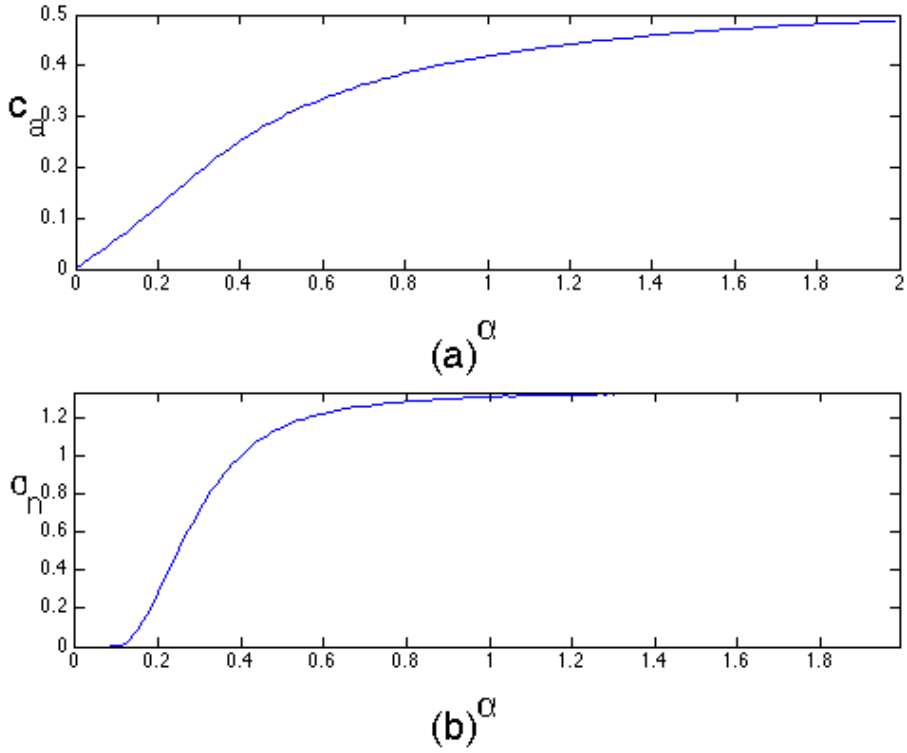

Figure 1.3: Variation dependence on α: (a) Coefficient of Variation of Area dependence on α; (b) Variation of Number of Neighbor dependence on α

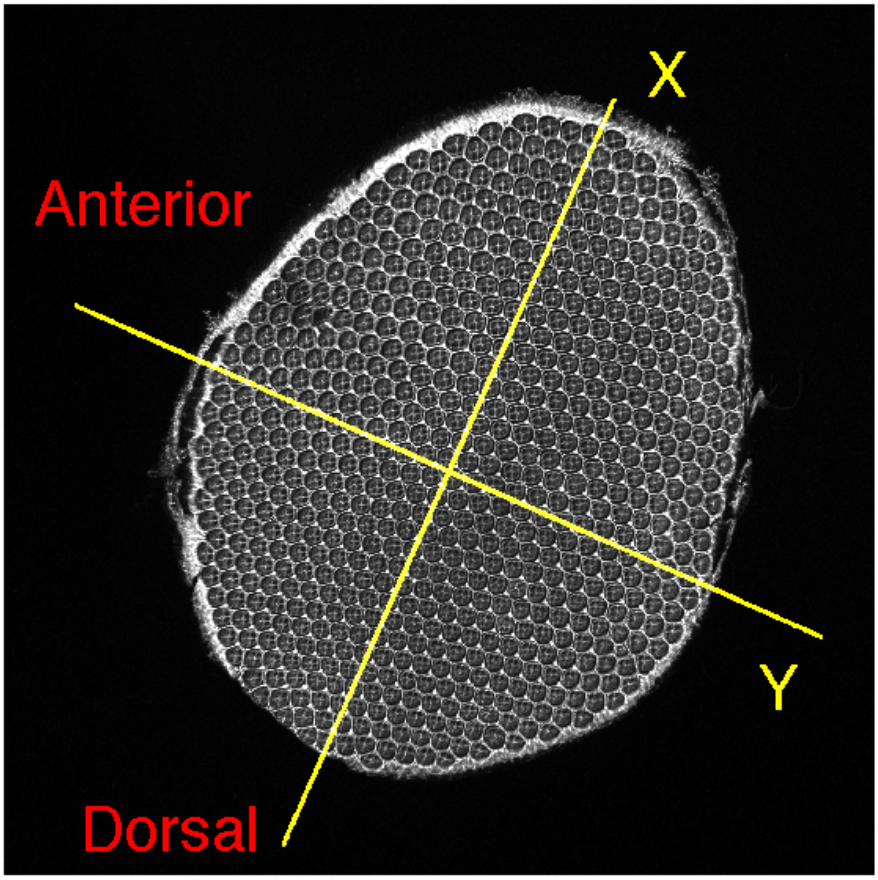

Figure 1.4: Experiment image of Drosophila entire retina, with long axis labeled X, short axis labeled Y and orientations relative to the fly's body indicated

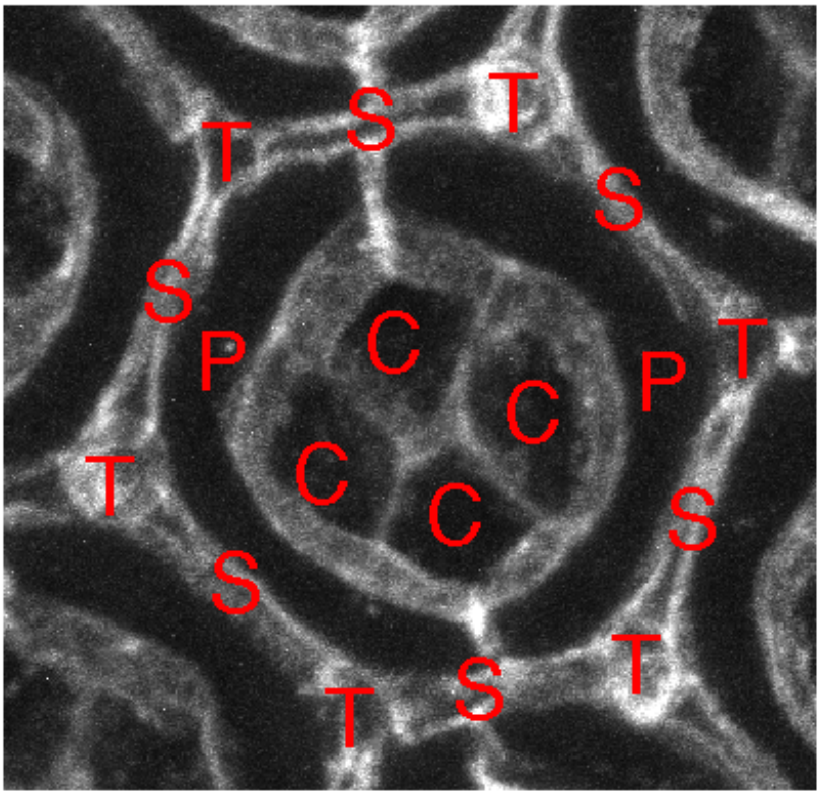

Figure 1.5: Individual ommatidium, with labels C (cone cells), P (primary pigment cells), S (secondary cells) and T (tertiary cells)

1.5 Tables

Table 1.1: Overall Coefficient of Variation of Area

Sample	Wild_a	Wild_b	Fz_a	Fz_e
$c_{A,tot}$	0.1146	0.1096	0.1762	0.1917
$\frac{c_{A,tot} - c_{A,critical}}{c_{A,critical}}\%$	64%	56%	152%	174%

CHAPTER 2: EXPERIMENT METHODS

2.1 Experimental Images Information

The experimental images are provided by Dr. Justin Cassidy of Professor Richard Carthew's research group at Northwestern University. The images contain pupal retinas of wild type and frizzled null genotypes at 48 to 54 hours after puparium formation at 25 degree Celsius. Wild samples have full genotype w1118. The Frizzled Null mutation samples have genotype w1118 with Fz (chromosomal deficiency GF3b) and Fz (chromosomal deficiency Exel6122). The pupal retinas have been antibody stained for the protein 'Discs Large' (Dlg) and pseudocolored in greyscale. This protein marks the cell's septate junction and is positioned basolaterally to the adherens junction. The samples were taken under 20X magnification; the confocal slices have 1.5 micron optical thickness. Figure 1.4 shows a projection of the confocal stack. The detailed information for sample size and image resolution is listed in Table 2.1.

2.2 Image Segmentation

Image segmentation has always been a crucial part of biological information analysis. Very often it requires a careful combination of different algorithm to extract proper information from the experimental image. The main algorithm used here is a type of watershed algorithm. Other morphological operation and manual correction of the image serve as a supplement.

2.2.1 Watershed Segmentation

Watershed segmentation is often used in image segmentation. It separates the image into regions based on local minima of intensity, without threshold parameters. In a grey scale image, for instance, pixel intensity can be seen as height. A local minimum will be dark pixels surrounded by light pixels. The notion was first introduced by Beucher and Lantuejoul in contour detection [17]. They make an excellent analogy of the segmentation with geographical notions. An image is regarded as a topographic surface. A drop of water on one point of the surface will flow to the local minimum. The set of all points that flow to the same minimum is considered as one region. As the water level rises, two nearby region will overlap on several points, see Figure 2.1. Those points form watershed lines that separate each local minimum. As mentioned before, the antibody stained protein Dlg is positioned at the adhesion junction on cell-cell interfaces, which show strong signal in the image. So the watershed lines will represent the ommaditia outlines. Watershed segmentation has many different algorithms. We use the Matlab build-in function 'watershed', which is based on the Fernand Meyer algorithm [18].

2.2.2 Oversegmentation

Oversegmentation is a common problem for watershed algorithm. Raw images, due to different reasons, almost inevitably have more local minima than those the user is interested in detecting. Procedures for pre-processing of the image are case specific. In our case, the slices in the image stacks show the cross section of the retina at different depths. Because the stained protein Dlg is positioned at different depths for different cell-cell interface, some images show strong signal in the ommatidium perimeter and some show strong signal in the central cone-cell interfaces.

Imposing minima in the images is a reasonable approach for pre-processing. For those images that show strong signal in the perimeter of ommatidium, if one and only one local minimum is imposed inside that ommatidium, the watershed algorithm will detect the perimeter as the watershed line. In our case, we use the built-in Matlab function 'imimposemin' to impose local minima at the position of markers [19]. The function is based on the fast hybrid grayscale reconstruction algorithm. A combination of different practices is used in order to find one and only one marker inside each ommatidium. We can use cone cells, which are positioned at the center of each ommatidium, as a marker. For images with a strong signal of cone cells, certain threshold values give a binary image of cone cells with noise. The noise can be eliminated with an area filter, which excludes objects smaller than a critical area, or by an image morphology process like 'image open', which excludes objects smaller than a given shape. Hand correction makes for a good complement. For cases like patches of missing signal intensity, hand correction is a fast and valid method to label the ommatidium center or to complete an ommatidum boundary.

With one and only one marker inside each ommatidum and the images of ommatidum frame cells, the watershed algorithm gives a skeletonized frame. The illustration for a typical process is displayed in Figure 2.2. Certain geometrical and topological properties can be extracted.

2.3 Geometrical and Topological Properties Extraction

The watershed line identifies every ommatidium with a closed contour of single-pixel width. Properties like area, orientation, neighbor relations and centroid can be extracted from the image.

Frist, we define the retina images' long axis as X axis and short axis as Y axis for consistency. Considering the overall eye orientation, the long axis is parallel to dorsal-ventral axis and the short axis is parallel to anterior-posterior axis. The long and short axes' direction of the image are defined as the long and short axes' direction of the ellipse that has the same second moments as the image. The centroid of the retina is defined as the coordinate origin.

Then, every ommatidium is assigned with a unique ID. This is done by Matlab function 'bwlabel'. The segmentation image is transformed into a label matrix of the same size. Every ommatidium will have a unique non-zero value and the watershed line will have a value zero for the corresponding elements in the label matrix. The area and centroid for each ommatidium can be calculated with the Matlab function 'regionprops'.

The neighbor relation can be obtained by many methods. A simple and straightforward method is used, as shown in Figure 2.3. For every element in the label matrix that has a value zero, its eight connected neighbor elements are examined. Because the watershed line is of single-pixel width, these neighbor elements can have value zero, which also belongs to the watershed line, or non-zero value, which belongs to the corresponding ommatidium. Further, if the element has three different non-zero neighbors, then it is a vertex, while if the element has two different non-zero neighbors, it is an edge. In this way, the number of neighbors and the neighbor relation for each ommatidium can be obtained.

2.4 Figures

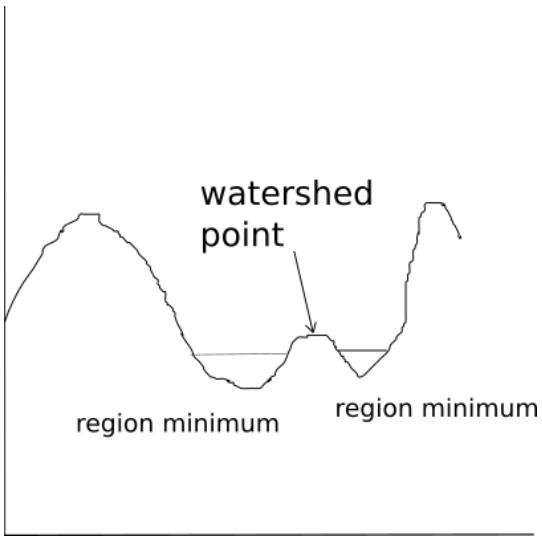

Figure 2.1: Illustration for Watershed algorithm

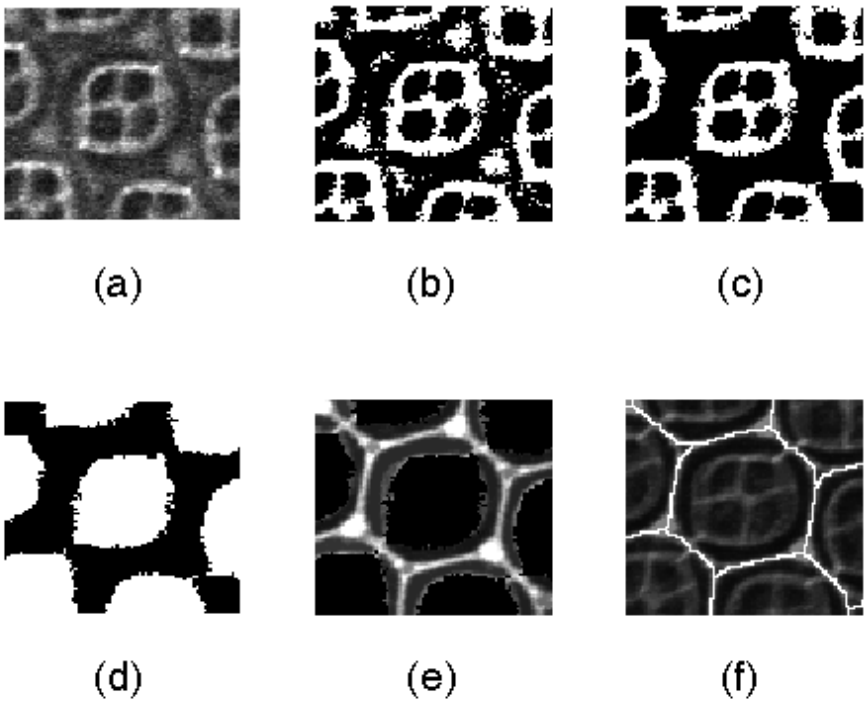

Figure 2.2: Illustration for segmentation process, (a) a slice of image with strong cone cells signal (b) binary image for certain threshold (c) filtered image (d) use cone cells postitions as markers for minima (e) minima imposed image (f) watershed line

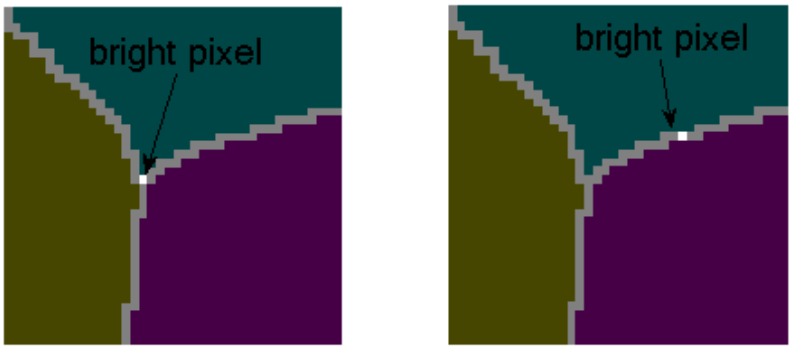

Figure 2.3: (a) Bright pixel has three different color pixels in eight-connected neighbor (b) Bright pixel has two different color pixels in eight-connected neighbor

2.5 Tables

Table 2.1: Sample size and corresponding image resolution

Sample Name	Width (μm)	Height (μm)	Depth (μm)	Resolution (pixel per μm)	Voxel (μm^3)
Wild_a	636.40	636.40	6.79	3.218	$0.31\times0.31\times0.34$
Wild_b	636.40	636.40	2.04	3.218	$0.31\times0.31\times0.34$
Fz_a	636.40	636.40	2.58	3.218	$0.31\times0.31\times0.43$
Fz_e	636.40	636.40	2.61	3.218	$0.31\times0.31\times0.43$

CHAPTER 3: SIMULATION

3.1 Simulation Methods

To study the relation of systematical variation, statistical variation and total variation of area, a simulation of a Voronoi tiling with imposed variation is developed.

The simulation bases on and extends from Lucarini's model. The Voronoi tiling of points in an equal lateral triangular grid gives a regular hexagon pattern. The lateral length $r_{x,ij}$ and $r_{y,ij}$ control the area of Voronoi cells, see Figure 3.1. By changing $r_{x,ij}$ and $r_{y,ij}$, we can manually impose a systematical variation of area. For simplicity, a linear area gradient is imposed. The corresponding parameter changes are $r_{x,(i+1)j_0} = r_{x,ij_0} - \beta_x r_0$ and $r_{y,i_0(j+1)} = r_{y,i_0 j} - \beta_y r_0$, where r_0 is a constant, i_0, j_0 are fixed indices and β_x or β_y are the parameters controlling the gradient in the x or y direction, respectively. The indices change as $i = 0,1,\cdots 59$ and $j = -14,-13,\cdots 13,14$. For a fixed i_0, $r_{x,\,i_0 j}$ remain constant for all j and $r_{y,i_0 0} = r_{x,\,i_0 0}$. The grid starts with $r_{x,00} = r_0$.

Then, statistical variation is imposed by adding Gaussian noise to the spatial position of grid points with magnitude $\varepsilon = \alpha(r_{x,ij} r_{y,ij})^{0.5}$. Different combinations of β_x and β_y determine the fixed area gradient vector direction and the fixed degree of systematical variation of area.

To get the Voronoi tiling of the grid points, we use the Matlab function 'voronoin'. It returns Voronoi edges V, with the positions of starting and ending points of each edge, and Voronoi cells C, with indices of all Voronoi edges belonging to each cell. The Voronoi cells on the boundary are excluded. The final tessellation is roughly 17×42 and 700 Voronoi cells, and thus similar to the sample size in experiment images. One sample Voronoi Tiling is displayed in Figure 3.2.

The geometrical and topological properties for each Voronoi cell are calculated. The number of neighbors is the number of Voronoi edges belonging to the Voronoi cell. If the spatial coordinates of Voronoi vertices for one Voronoi cell in the sequence is $\{x_1, x_2, \cdots x_n\}$ and $\{y_1, y_2, \cdots y_n\}$, then the area for that cell is given by

$$A = \frac{1}{2}\sum_{i=1}^{n}(x_i y_{i+1} - x_{i+1} y_i) \tag{3.1}$$

with negative value if the vertices are in clockwise order and positive value if counterclockwise order. Then, the centroid is given by

$$\begin{cases} C_x = \dfrac{1}{6A}\displaystyle\sum_{i=1}^{n}(x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \\ C_y = \dfrac{1}{6A}\displaystyle\sum_{i=1}^{n}(y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \end{cases} \tag{3.2}$$

With centroid and area for each Voronoi cell calculated, a plane can be obtained by least square fitting a function $f(x, y)$ to the empirical values $f_0(C_x, C_y)$, see Figure 3.3. In general, the least square fitting procedure can be concluded as finding a set of parameters that makes the sum of squared residuals $\sum(f - f_0)^2$ a minimum. For a polynomial least square fitting, function $f(x, y)$ can be written as,

$$f = p_{00} + p_{10}x + p_{01}y + p_{20}x^2 + p_{11}xy + p_{02}y^2 + \cdots \tag{3.3}$$

where p_{ij} are parameters of $(i + j)$ degree. In a plane fitting, $p_{ij} = 0$ for $i + j > 1$. In a quadratic fitting, $p_{ij} = 0$ for $i + j > 2$. The corresponding $c_{A,sys}$ can be calculated. Together with the total coefficient of variation of area $c_{A,tot}$, we can have a derived coefficient of statistical variation of area $c'_{A,stat}$ from the square rule. With prior knowledge of the magnitude of imposed Gaussian noise, we should have an expected $c_{A,stat}$. By comparing $c'_{A,stat}$ and $c_{A,stat}$, we can check the accuracy of the independence of statistical and systematic variations.

3.2 Figures

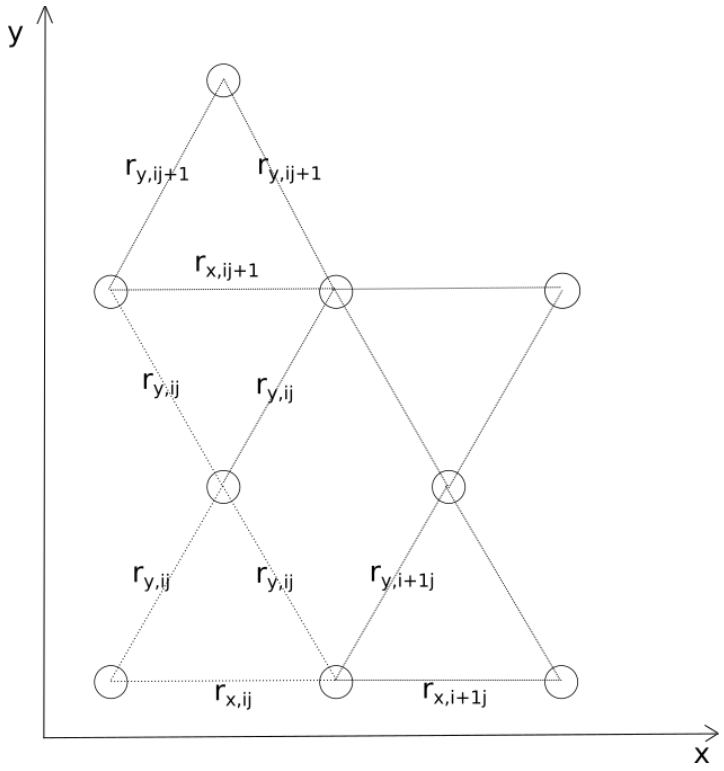

Figure 3.1: Grid Lateral Length

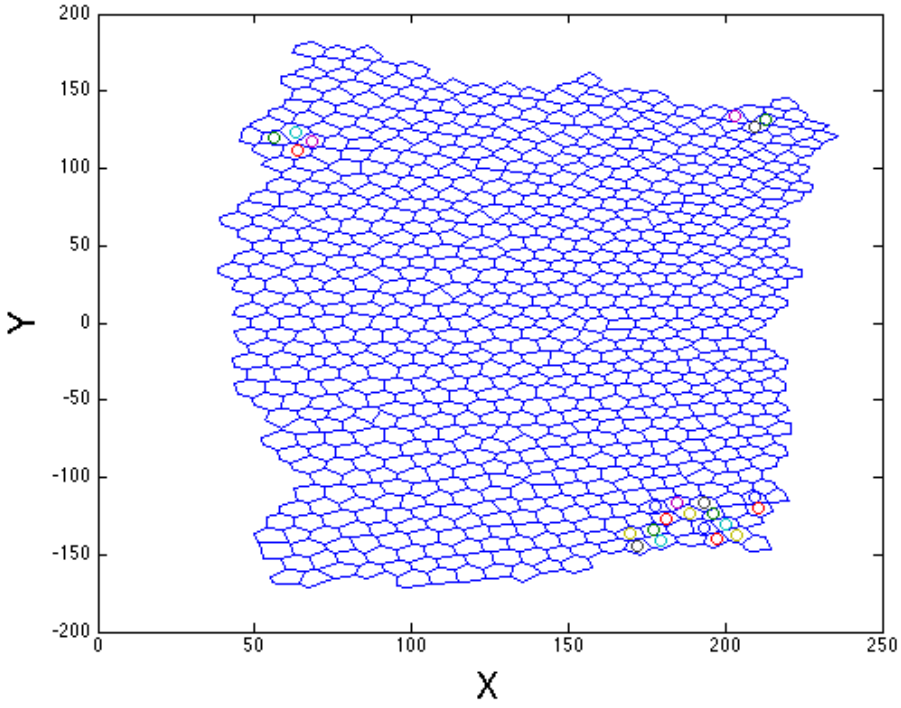

Figure 3.2: Voronoi Tiling with $c_{A,sys} = 0.1533$ and $c_{A,stat} = 0.0896$, Defects Labeled with Color Circles

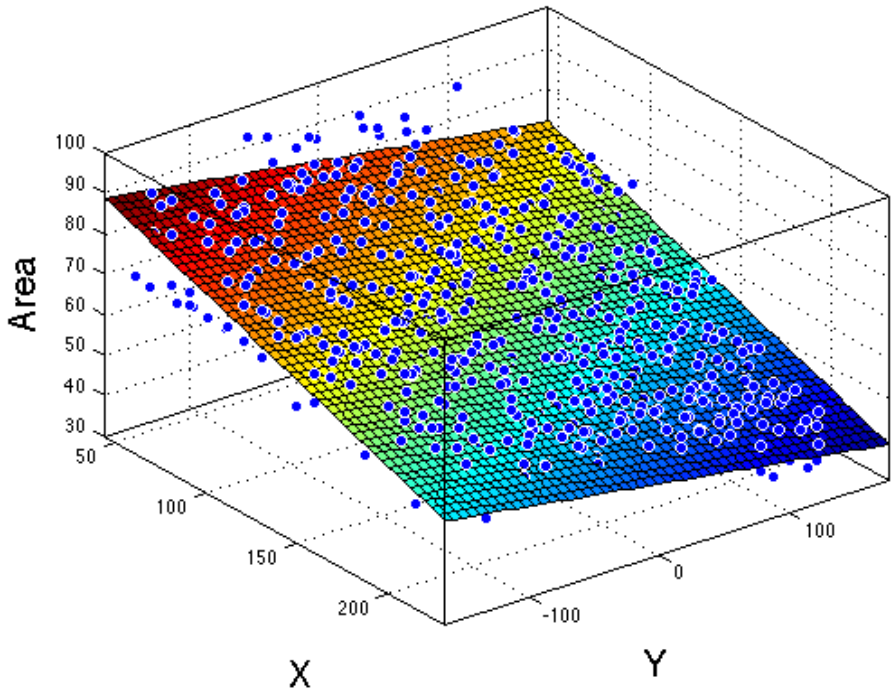

Figure 3.3: Plane Fitting for Voronoi Tiling in Figure 3.2

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Statistical and Systematical Variation Independence Verification

The simulation has linear area change in both x and y direction. Thus a planar fitting should yield an unambiguous result. Different combinations of gradient parameters β_x and β_y have different magnitude of systematical variation and direction of gradient vector. The accuracy of systematical variation is based on the quality of the fitting. Even with the prior knowledge of the direction and the magnitude of area changes, error in fitting is inevitable. But we should expect them to have minor impact on the result.

For simplicity, two cases are tried. In the first case, the gradient parameter $\beta_x = \beta_y = \beta$ and changes from 0 to 0.001 with interval 0.00005 and 0.001 to 0.015 with interval 0.001. For each gradient parameter value, the Gaussian noise parameter changes from 0 to 0.5 with interval 0.01. The fitting shows that the direction of the gradient vector is 16 degree with respect to x-axis and the coefficient of systematical variation of area $c_{A,sys}$ ranges from 0 to 0.3. In the second case, gradient parameter $\beta_y = 0$ and $\beta_x = \beta$, the rest is the same as the first case. Now the direction of the gradient vector is parallel to the x-axis.

For each β and α, $c_{A,sys}$ and $c_{A,tot}$ are calculated. Then we compare the difference between $c'_{A,stat}$ from square rule and $c_{A,stat}$ from the imposed α. First, when $\beta = 0$, the α - $c'_{A,stat}$ plot is a perfect match with α-$c_{A,stat}$ plot. Then the difference between $c'_{A,stat}$ and $c_{A,stat}$ is studied for different gradient. For both cases, the absolute difference $\delta = c'_{A,stat} - c_{A,stat}$ increases as α or β increases, see Figure 4.1 and Figure 4.2. The relative difference $\%\delta = (c'_{A,stat} - c_{A,stat})/c_{A,stat}$ gives a better understanding, see Figure 4.3 and Figure 4.4. For both cases, the relative

difference is large when $c_{A,stat}$ is considerably smaller than $c_{A,sys}$. In such situation, the error in $c_{A,sys}$ due to fitting error is of similar order as $c_{A,stat}$. For the same reason, the direction of imposed area gradient direction will be a significant effect on $\%\delta$. But in the parameter regions of interest, which for wild-type images are $c_{A,sys} \in [0.08, 0.10]$ and $\alpha \in [0.10, 0.14]$ and for mutant images are $c_{A,sys} \in [0.14, 0.16]$ and $\alpha \in [0.17, 0.21]$, the largest relative difference is around 3.4% for wild-type and 6.7% for the mutant parameter range. It will be sufficient to say that the square rule is an accurate way of calculating the coefficient of statistical variation of area $c_{A,stat}$.

4.2 Experiment Result

Based on square rule, $c_{A,stat}$ for different images are listed in Table 4.1. It turns out that $c_{A,stat}$ for wild-type images are around the critical value that would allow first defects to appear while $c_{A,stat}$ values for mutant images are far beyond the critical value, necessitating defects. It shows that experiment data is consistent with the theory of symmetry breaking in Voronoi tiling. Moreover, in Frizzled mutation sample-a, quadratic fitting gives better result than plane fitting, see Table 4.1. This indicates that drosophila retina can have different trends of ommaditium area change and a proper fitting surface is important in understanding $c_{A,stat}$ and $c_{A,sys}$. There is, however, another factor that biases statistical variation towards larger values: finite image resolution. The following section estimates this effect.

4.3 Image Processing Error Estimation

Estimating the error of the watershed algorithm is also important. A possible estimation is proposed here. The error of the watershed Algorithm is related to the width of frame cells in the image, which is roughly 2 to 3 pixel-widths. Now if we

start with an imaginary completely regular hexagon tessellation, then no statistical variation of area exists. If each edge can fluctuate among 3 pixel-width with equal probability, then one of its touching cells has the same probability in losing area δ, remaining the same area or gaining area δ. The area losing and gaining corresponds to the area gained and lost from the other touching cell. In simulation, δ is the same as the edge length in pixel. An ommatidium with average edge length 29 pixels (representative of experimental-image ommatidia) has roughly 2185 pixel in area. So $\delta \approx 29$.

The simulation shows that the watershed induced coefficient of variation of area is roughly 0.0262. If we assume the independence of the watershed error and statistical variation of area (certainly a good assumption in this case), then $c_{A,stat}$ for experiment is further reduced, although only by a small amount. The result is displayed in Table 4.2; it shows still the same qualitative result of defect-free lattices being possible for wild-type area variations, but not for frizzled mutants.

4.4 Coefficient of Variation of Number of Neighbor

Beyond the qualitative appearance of defects, can we understand or predict the number of defects that should occur in the mutants? Using quadratic fitting and eliminating watershed error, both Frizzled mutation samples have $c_n \approx 0.08$, estimated $c_{A,stat} \approx 0.11$ and $c_{A,tot} \approx 0.18$. The $c_{A,stat}$ - c_n plot generated from simulation with no systematical variation shows that c_n in experiment are not consistent with the derived $c_{A,stat}$, see Figure 4.5. It suggests that either the systematical variation will contribute to defects generation or other mechanisms are involved to reason the derived $c_{A,stat}$ and empirical c_n.

One possible argument is that systematical variation of area contributes to defects generation. Figure 4.6 shows $c_{A,stat}$-c_n plot with no systematical variation and plot with $c_{A,sys} \in [0.13,0.16]$, which is the $c_{A,sys}$ range calculated from quadratic fitting of empirical data. In Figure 4.6, for small $c_{A,stat}$, $c_n$with no systematical variation is clearly smaller than c_n with $c_{A,sys}$ of magnitude derived from empirical data. Meanwhile, the legitimacy of neighbor relation derived from image segmentation remains questioning. The bias in resolving four-way junctions has influence on c_n. The segmented image will always identify an edge in the location of four-way junction, see Figure 4.7. But the flip of that edge can change a local 5, 5, 7, 7 defects combination into no defect at all. It might significantly reduce c_n found in experiment. Miklius et al. suggest that in drosophila wing epithelium, the four-way junctions tend to resolve in a way that gives the narrowest number of neighbor distribution [20]. If we assume that four-way junctions in retina resolved in a similar manner, we will obtain a new segmentation with fewer defects, see Figure 4.8. For instance, the new segmentation for Fz_e sample has $c_n' = 0.0615$, which is significantly smaller than the old $c_n = 0.0789$. Moreover, c_n' fits with $c_{A,stat}$-c_n plot for $c_{A,stat}$ of Fz_e sample, see Figure 4.6. In this way, the inconsistency of $c_{A,stat}$ calculated and c_n observed is solved. Further image analysis may confirm this argument.

Another postulation is that systematical variation of area will not contribute to the variation of number of neighbor, or in other words, the variation of number of neighbor is caused only by statistical variation of area, then modification of the model is necessary. A possible argument is proposed here. While the magnitude of statistical variation of area is uniform in wild sample, different region in mutant sample can have different magnitude of statistical variation of area. For simplicity, further

assumptions are made. First, the mutant samples have only two regions with two different magnitude of statistical variation of area, indicated by α_1 and α_2. Second, both regions have same mean area μ_A. If we assume region one has area portion $\gamma(0 < \gamma < 1)$, then region two has area portion $1 - \gamma$. Consequently the following set of equations has to be fulfilled,

$$\begin{cases} c_{A,1}^2 \cdot \gamma + c_{A,2}^2 \cdot (1-\gamma) = c_{A,stat}^2 \\ c_{n,1}^2 \cdot \gamma + c_{n,2}^2 \cdot (1-\gamma) = c_n^2 \end{cases}$$

where indices 1 and 2 indicate which region the parameter belongs to.

Some sets of parameters α_1, α_2 and γ are found to have $c_{A,stat}$ and c_n within the experiment domain ($c_{A,stat} \approx 0.1133$ and $c_n \approx 0.0789$), see Table 4.3. The result suggests that one large portion of retina is very ordered while the rest is extremely disorder.

4.5 Morphological Orientation of Systematic Variation

A plane fitting gives a gradient vector with a certain direction. The directions for the experimental images show some consistency. For instance, the gradient vector in plane fitting of wild-type images points from top dorsal and front anterior part of the retina toward the ventral and posterior part. This consistency may suggest that the region of retina where the ommatidia are generally larger is directly related to their position during the dynamics of morphogenesis. Similarly, in quadratic fitting, the eigenvector of the Hessian matrix corresponding to its larger eigenvalue points in the direction of largest curvature. In the samples, this direction is 42 degree with respect to long axis of retina in both wild images and 15 degree in Fz-a mutant image and 20 degree in Fz-e mutant image. The direction of curvature shows consistency in wild-type image. Further study with more sample analysis will firm up the existence of

these systematic directions of area variation and, if they are confirmed, will formulate hypotheses on the mechanisms guiding these features.

4.6 Figures

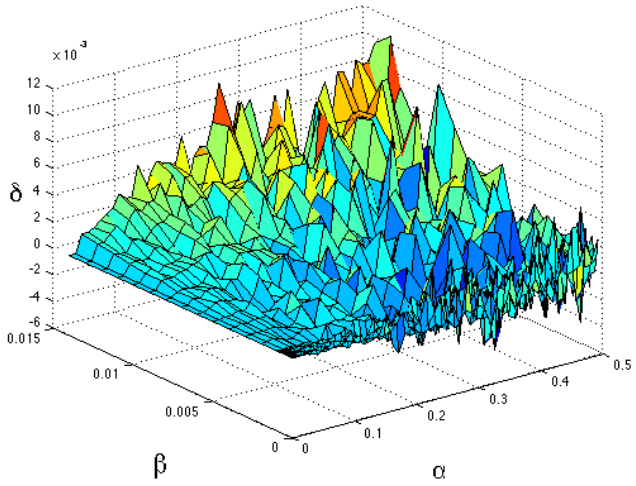

Figure 4.1: The absolute difference between $c'_{A,stat}$ and $c_{A,stat}$ with gradient vector direction parallel to x, $\beta_x = \beta, \beta_y = 0$

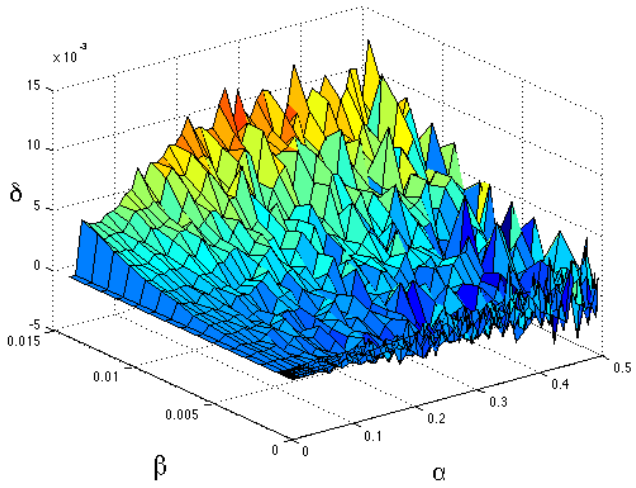

Figure 4.2: The absolute difference between $c'_{A,stat}$ and $c_{A,stat}$ with gradient vector direction parallel to x, $\beta_x = \beta_y = \beta$

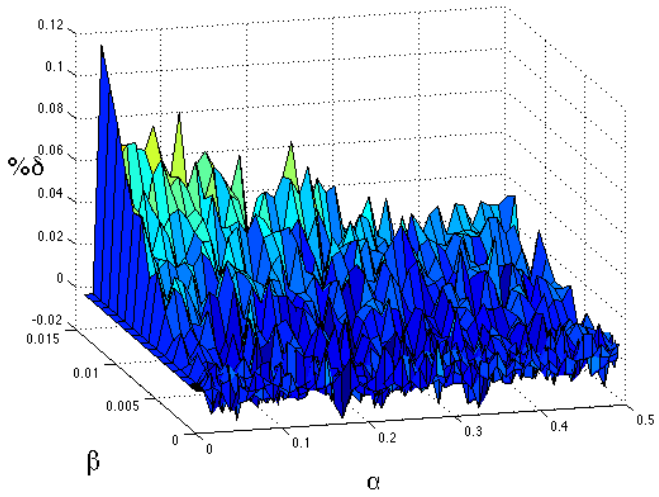

Figure 4.3: The relative difference between $c'_{A,stat}$ and $c_{A,stat}$ with gradient vector direction parallel to x, $\beta_x = \beta, \beta_y = 0$

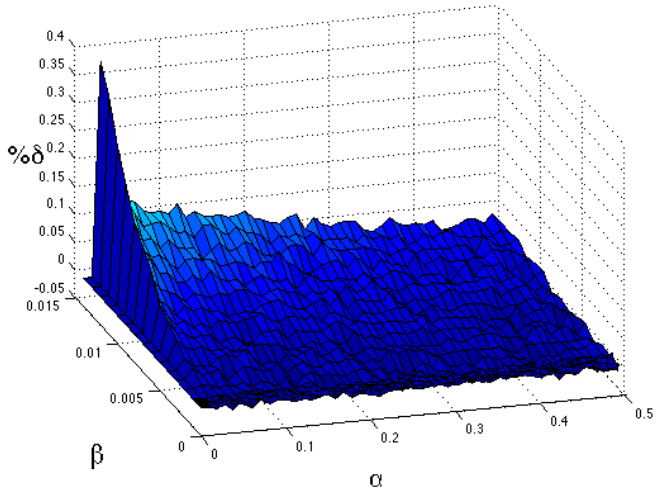

Figure 4.4: The relative difference between $c'_{A,stat}$ and $c_{A,stat}$ with gradient vector direction parallel to x, $\beta_x = \beta_y = \beta$

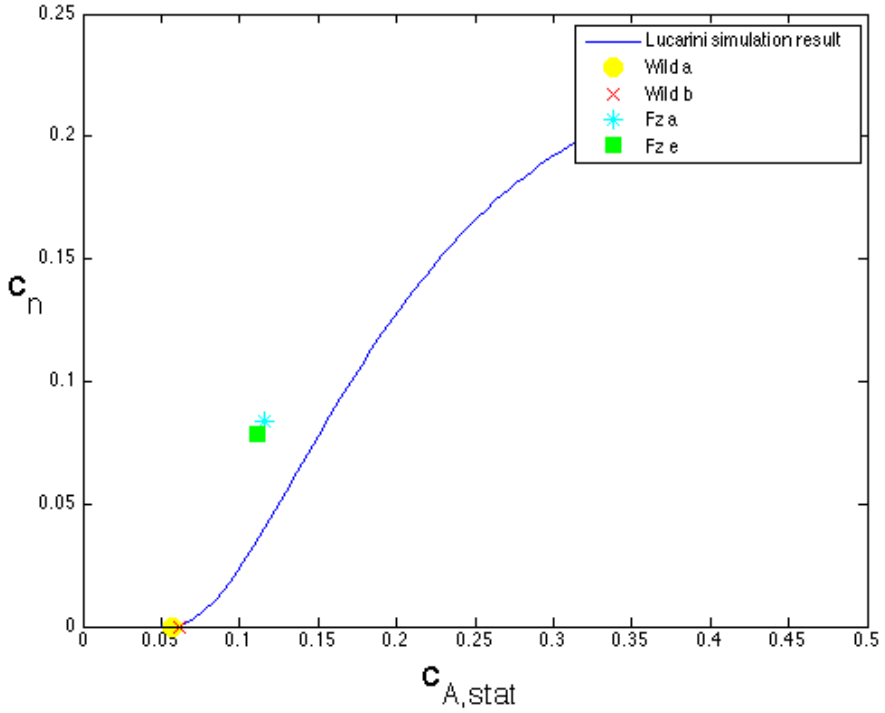

Figure 4.5: The $c_{A,stat}$-c_n plot with no systematical variation from Lucarini simulation result, the markers represent derived $c_{A,stat}$ from quadratic fitting and empirical c_n, with yellow dot for Wild_a, red cross for Wild_b, cyan aserisk for Fz_a and green square for Fz_e

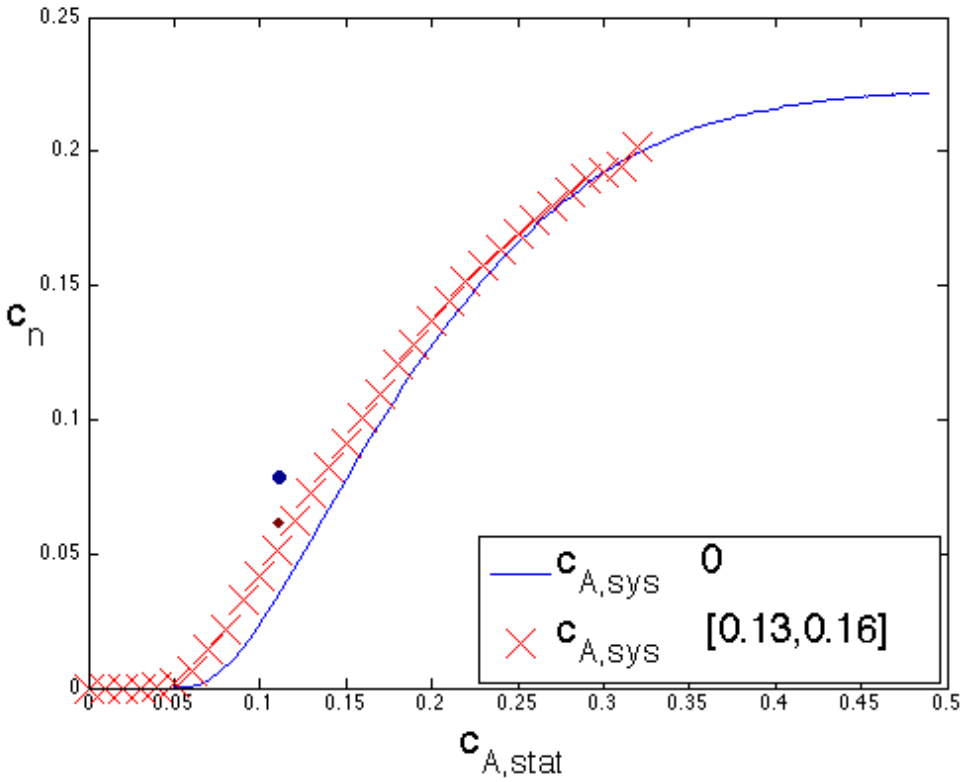

Figure 4.6: The $c_{A,stat}$-c_n plot with no systematical variation and $c_{A,sys} \in [0.13,0.16]$. Blue circle represents c_n in original segmentation; red diamond represents c_n in four-way junction resolved segmentation.

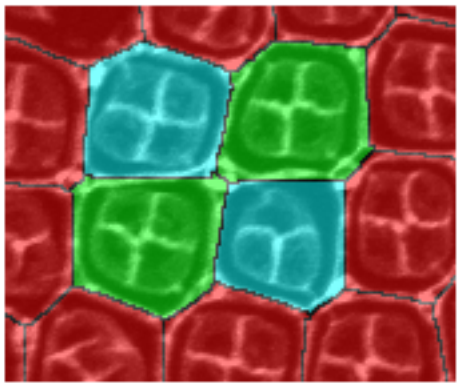

Figure 4.7: Four-way junction, red represents six-neighbor cells, green represents seven-neighbor cells, and cyan represents five-neighbor cells

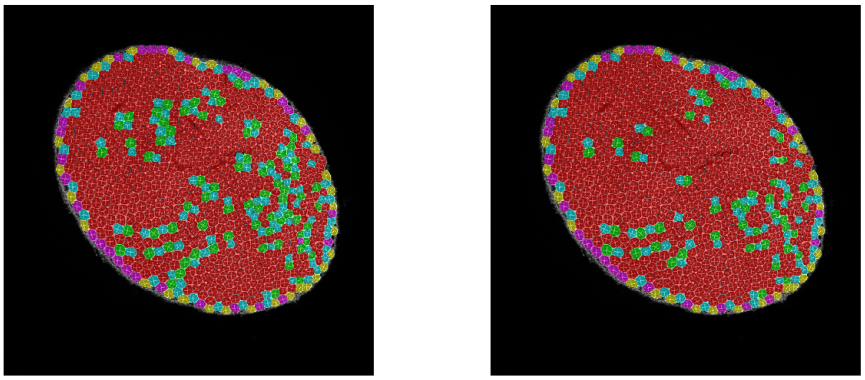

Figure 4.8: (a) Original segmentation (b) Four-way junctions resolved segmentation; yellow for 3-sided, magenta for 4-sided, cyan for 5-sided, red for 6-sided and green for 7-sided cells.

4.7 Tables

Table 4.1: Experiment fitting result

Sample	$c_{A,tot}$	Plane fit $c'_{A,stat}$	Corresponding $c_{A,sys}$	Quadratic fit $c'_{A,stat}$	Corresponding $c_{A,sys}$
Wild_a	0.1146	0.0698	0.0909	0.0625	0.0961
Wild_b	0.1096	0.0763	0.0786	0.0671	0.0866
Fz_a	0.1762	0.1512	0.0905	0.1189	0.1300
Fz_e	0.1917	0.1220	0.1479	0.1138	0.1543

Table 4.2: Watershed error reduction from quadratic fit

Sample	$c_{A,stat}$ without watershed error	$c_{A,stat}$ with watershed error
Wild_a	0.0567	0.0625
Wild_b	0.0618	0.0671
Fz_a	0.1160	0.1189
Fz_e	0.1107	0.1138

Table 4.3: Parameter sets for different region in retina

α_1	α_2	γ	$c_{A,stat}$	c_n
0.04	0.38	0.77	0.1172	0.0766
0.03	0.39	0.78	0.1165	0.0766
0.03	0.38	0.77	0.1163	0.0766

CHAPTER 5: CONCLUSION

The study of geometrical and topological properties of ommatidial pattern in the Drosophila eye provides insight into the mechanism of genetic mutations and their influence on the planar cell polarity pathway. Our study analyzed ommatidial areas systematically for the first time. The analysis unexpectedly found repeatable systematic area gradients in wild type Drosophila eyes. Such a variation may suggest that the region of retina where the ommatidia are generally larger is directly related to their position during the dynamics of morphogenesis. Meanwhile, previous studies on general 2D tilings, based on models with no systematic variability, found an area variation restriction for hexagonal tessellation [8] [9], which the empirical area variation is inconsistent with. We separated the systematic variability from the statistical variability and showed that the statistical variability does explain the non-existence of defects in wild type and the existence of defects in Fz mutants. This may suggest that the Fz mutant ommatidia lose large-scale orientation not directly because of loss of planar cell polarity, but due to the existence of defects, necessitated by a loss of sufficient area uniformity. Further studies on other mutants will confirm or modify these conclusions.

REFERENCE

[1] Koonin EV, "Orthologs, Paralogs, and Evolutionary Genomics," *Annual Review of Genetics*, vol. 39, pp. 309-338, December 2005.

[2] Richard W Carthew, "Pattern Formation in the Drosophila Eye," *Current Opinion in Genetics & Development*, vol. 17, no. 4, pp. 309-313, August 2007.

[3] Takashi Hayashi and Richard W Carthew, "Surface Mechanics Mediate Pattern Formation in the Developing Retina," *Nature*, vol. 431, pp. 647-652, October 2004.

[4] Anne-Kathrin Classen, Kurt I. Anderson, Eric Marois, and Suzanne Eaton, "Hexagonal Packing of Drosophila Wing Epithelial Cells by the Planar Cell Polarity Pathway," *Developmental Cell*, vol. 9, no. 6, pp. 805-817, December 2005.

[5] Susmit Kumar and Stewart K, Kurtz, "Properties of a Two-Dimensional Poisson-Voronoi Tesselation: a Monte-Carlo Study," *Materials Characterization*, vol. 31, no. 1, pp. 55-68, July 1993.

[6] Frederic T Lewis, "The Correlation between Cell Division and the Shapes and Sizes of Prismatic Cells in the Epidermis of Cucumis," *The Anatomical Record*, vol. 38, pp. 341-376, 1928.

[7] J Lemaitre et al., "Arrgangement of Cells in Voronoi Tesselations of Monosize Packing of Discs," *Philosophical Magazine Part B*, vol. 67, no. 3, pp. 347-362, 1993.

[8] Matthew P Miklius and Sascha Hilgenfeldt, "Analytical Results for Size-Topology Correlations in 2D Disk and Cellular Packings," *Physical Review Letters*, vol. 108, no. 1, p. 015502, January 2012.

[9] Valerio Lucarini, "From Symmetry Breaking to Poisson Point Process in 2D Voronoi Tessellaions: the Generic Nature of Hexagons," *Journal of Statistical Physics*, vol. 130, no. 6, pp. 1047-1062, March 2008.

[10] Eberhard Bodenschatz, Werner Pesch, and Guenter Ahlers, "Recent Developments in Rayleigh-Benard Convection," *Annual Review of Fluid Mechaniscs*, vol. 32, pp. 709-778, January 2000.

[11] Ian M Gemp, Richard W Carthew, and Sascha Hilgenfeldt, "Cadherin-Dependent Cell Morphology in an Epithelium: Constructing a Quantitiative Dynamical Model," *PLOS Computational Biology*, vol. 7, no. 7, July 2011.

[12] Richard W Carthew, "Adhesion Proteins and the Control of Cell Shape," *Current Opinion in Genetics & Development*, vol. 15, no. 4, pp. 358-363, August 2005.

[13] Sascha Hilgenfeldt, Sinem Erisken, and Richard W Carthew, "Physical Modeling of Cell Geometric Order in an Epithelial Tissue," *Proceddings of the National Academy of Sciences*, vol. 105, no. 3, pp. 907-911, January 2008.

[14] Jos Kafer, Takashi Hayashi, Athanasius F.M Maree, and Richard W Carthew, "Cell Adhesion and Cortex Contractility Determine Cell Patterning in the Drosophila Retina," *Proceedings of the National Academy of Sciences*, vol. 104, no. 47, pp. 18549-18554, Novermber 207.

[15] Christopher G Winter et al., "Drosophila Rho-Associated Kinase (Drok) Links Frizzled-Mediated Planar Cell Polarity Signaling to the Actin Cytoskeleton," *Cell*, vol. 105, pp. 81-91, April 2001.

[16] Michel Gho and Francois Schweisguth, "Frizzled Signalling Controls Orientation of Asymmetric Sense Organ Precursor Cell Divisions in Drosophila," *Nature*, vol. 393, pp. 178-181, May 1998.

[17] S Beucher and C Lantuejoul, "Use of Watershed in Contour Detection," in *International Workshop on Image Processing: Real-Time Edge and Motion Detection/Estimation*, Rennes, France, 1979.

[18] Fernand Meyer, "Topographic Distance and Watershed Lines," *Signal Processing*, vol. 38, no. 1, pp. 113-125, July 1994.

[19] Luc Vincent, "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms," *IEEE Transactions on Image Processing*, vol. 2, no. 2, pp. 176-201, April 1993.

[20] Matthew P Miklius and Sascha Hilgenfeldt, "Epithelial Tissue Statistics: Eliminating Bias Reveals Morphological and Morphogenetic Features," *The European Physical Journal E*, vol. 34, no. 50, May 211.